

Reference  
A-99ED0347

OH-522A  
Print Art 1 P-388  
(Reference 1)

⑨ 日本国特許庁 (JP)

⑩ 特許出願公開

⑫ 公開特許公報 (A) 昭63-46498

⑤Int.Cl.<sup>4</sup>  
G 10 L 5/00

識別記号  
厅内整理番号  
8622-5D

⑪公開 昭和63年(1988)2月27日

審査請求 未請求 発明の数 2 (全6頁)

④発明の名称 韻律制御方式

②特願 昭61-191394

②出願 昭61(1986)8月15日

優先権主張 ②昭61(1986)4月18日③日本(JP)①特願 昭61-89360

②発明者 酒寄哲也 東京都大田区中馬込1丁目3番6号 株式会社リコー内  
②発明者 佐々部昭一 東京都大田区中馬込1丁目3番6号 株式会社リコー内  
②発明者 北川博雄 東京都大田区中馬込1丁目3番6号 株式会社リコー内  
①出願人 株式会社リコー 東京都大田区中馬込1丁目3番6号  
④代理人 弁理士高野明近

明細書

1. 発明の名称

韻体制御方式

2. 特許請求の範囲

(1) 予め用意した音声素片のパラメータ系列を入力文字列に従って読み出し、結合規則によって接続し、韻律規則によって韻率を付加する音声規則合成装置において、種々のパラメータを多変量統計解析的に処理して最適な制御値を得て韻律を制御することを特徴とする韻律制御方式。

(2) 制御すべき音韻時間長に影響を与えると考えられる定性的パラメータを用いて数量化I類分析を行うことによって測定値との最小2乗誤差を最小にする予測モデルを作り、これによって制御することを特徴とする特許請求の範囲第(1)項記載の韻律制御方式。

(3) 制御すべき音韻時間長に影響を与えると考えられる定量的パラメータを用いて線形重回帰分析を行うことによって測定値との最小2乗誤差を最小にする予測モデルを作り、これによって制御

することを特徴とする特許請求の範囲第(1)項記載の韻律制御方式。

(4) 制御すべき音韻時間長に影響を与えると考えられる定量的パラメータから、主成分分析によってこれらの特性を最もよく表わす互いに無相関ないいくつかの新らしいパラメータを求め、これを音韻時間長予測モデルのパラメータとすることを特徴とする特許請求の範囲第(1)項記載の韻律制御方式。

(5) 本来定性的パラメータとして表わされる前後の音韻環境を音韻の周波数スペクトルを表す種々のパラメータの時間変化という定量的パラメータの形で、その他の定量パラメータに加えて分析することを特徴とする特許請求の範囲第(2)項又は第(3)項記載の韻律制御方式。

(6) 予め用意した音声素片のパラメータ系列を入力文字列に従って読み出し、結合規則によって接続し、韻律規則によって韻律を付加する音声規則合成装置において、種々のパラメータを多変量統計解析的に処理して最適な制御値を得てタイミ

ング点パターンを生成することを特徴とする韻律制御方式。

(7) タイミング点間隔に影響を与えると考えられる、定性的パラメータ（前後の音韻環境、発話単位のモーラ数、モーラ位置、アクセント情報、強意情報等）を用いて数量化し類分析を行なうことによって実測値との2乗誤差を最小にする予測モデルを作り、これによって予測したタイミング点間隔を連結してタイミング点パターンを生成することを特徴とする、特許請求の範囲第(6)項記載の韻律制御方式。

(8) タイミング点間隔に影響を与えると考えられる、定量的パラメータ（前後のタイミング点間隔、ピッチ変化、振幅変化等）を用いて線形重回帰分析を行なうことによって実測値との2乗誤差を最小にする予測モデルを作り、これによって予測したタイミング点間隔を連結してタイミング点パターンを生成することを特徴とする特許請求の範囲第(6)項記載の韻律制御方式。

(9) タイミング点間隔に影響を与えると考えら

れる、定量的パラメータ（前後のタイミング点間隔、ピッチ変化、振幅変化等）から、主成分分析によってこれらの特徴を最も良く表す互いに無相関ないくつかの新しいパラメータを求め、これをパラメータとする予測モデルによって求めたタイミング点間隔を連結してタイミング点パターンを生成することを特徴とする特許請求の範囲第(8)項記載の韻律制御方式。

(10) 本来定性的パラメータとして表される前後の音韻環境を、音韻の周波数スペクトルを表す種々のパラメータの時間変化という定量的パラメータの形でその他の定性的パラメータに加えて分析することを特徴とする特許請求範囲第(8)項又は第(9)項記載の韻律制御方式。

### 3. 発明の詳細な説明

#### 技術分野

本発明は、音声の規則合成における韻律制御方式、より詳細には、音韻時間長制御方式及びタイミング点パターン生成方式に関する。

#### 従来技術

音声合成において自然な韻律を付加するために、ピッチ、振幅、リズム等を制御する韻律制御規則が不可欠であるが、このうちリズムの制御には、音韻の継続時間長を制御する方式と、音節中の特定の部位に注目し、その部位（タイミング点）を時間軸上である規則に依って生成されるパターンに合わせて行く方式が存在する。このうちの音韻時間長制御方式としては、各音韻の前後及びその音韻自身の時間長の平均を予め求めておき、制御すべき音韻環境に応じてこれらを参照して加重平均をとる方式や、発話単位全体を表わすモーラ数や、モーラ位置からのみ時間長を求める方式や、1モーラの時間長を全て一定にする方式などが提案されている。しかし、これらの方程式は、

(1) 個々のパラメータを個別に扱っているため、全体として正しい時間長を得られる保証はない。

(2) どのようなパラメータをどの程度の重みで考慮するかを経験的に決めている。

等の欠点がある。

また、後者のタイミング点パターン生成方式に

は、タイミング点間隔を等間隔とする方式や、発話単位内モーラ数や、前後の音韻環境等の種々のパラメータの影響を考慮して修正を加えるなどの方式が提案されている。しかしこの方式には以下の問題点がある。

(3) 個々のパラメータを別々に取扱っているために全体として正しいタイミング点パターンを得られることが保証されない。

等の欠点がある。

#### 目的

本発明は、上述のごとき実情に鑑みてなされたもので、特に、音声の規則合成において自然な韻律を生成するように韻律を制御すること、更に詳細に説明すると、自然な韻律を生成するように音韻の時間長を制御すること、及び、韻律の自然性を高めるように合成音声のタイミング点パターンを生成することを目的としてなされたものである。

#### 構成

本発明は、上記目的を達成するために、本発明の第1の実施例は、

(1) 予め用意した音声断片のパラメータ系列を入力文字列に従って読み出し、結合規則によって接続し、韻律規則によって韻律を付加する音声規則合成装置において、種々のパラメータを多変量統計解析的に処理して最適な制御値を得て音韻時間長を制御すること。

(2) 制御すべき音韻時間長に影響を与えると考えられる定性的パラメータ（前後の音韻環境、発話単位のモーラ数、発話単位内のモーラ位置等）を用いて数量化I類分析（カテゴリカル重回帰分析）を行うことによって測定値との最小2乗誤差を最小にする予測モデルを作り、これによって制御すること。

(3) 制御すべき音韻時間長に影響を与えると考えられる定性的パラメータ（前後の音韻時間長、前後のピッチ変化、パワー変化等）を用いて様形重回帰分析を行うことによって測定値との最小2乗誤差を最小にする予測モデルを作り、これによって制御すること。

(4) 制御すべき音韻時間長に影響を与えると考

えられる定性的パラメータ（前後の音韻時間長、前後のピッチ変化、パワー変化等）から主成分分析によってこれらの特性を最もよく表わす互いに無相関ないくつかの新しいパラメータを求め、これを音韻時間長予測モデルのパラメータとすること。

(5) 本来定性的パラメータとして表わされる前後の音韻環境を音韻の周波数スペクトルを表す種々のパラメータの時間変化という定延的パラメータの形で、その他の定量パラメータに加えて分析すること、を特徴としたものであり。

また、本発明の他の実施例は、

(6) 予め用意した音声断片のパラメータ系列を入力文字列に従って読み出し、結合規則によって接続し、韻律規則によって韻律を付加する音声規則合成装置において、種々のパラメータを多変量統計解析的に処理して最適な制御値を得てタイミング点パターンを生成すること。

(7) タイミング点間隔に影響を与えると考えられる、定性的パラメータ（前後の音韻環境、発話

単位のモーラ数、モーラ位置、アクセント情報、強意情報等）を用いて数量化I類分析を行なうことによって実測値との2乗誤差を最小にする予測モデルを作り、これによって予測したタイミング点間隔を連結してタイミング点パターンを生成すること。

(8) タイミング点間隔に影響を与えると考えられる、定性的パラメータ（前後のタイミング点間隔、ピッチ変化、振幅変化等）を用いて様形重回帰分析を行なうことによって実測値との2乗誤差を最小にする予測モデルを作り、これによって予測したタイミング点間隔を連結してタイミング点パターンを生成すること。

(9) タイミング点間隔に影響を与えると考えられる、定性的パラメータ（前後のタイミング点間隔、ピッチ変化、振幅変化等）から、主成分分析によってこれらの特徴を最も良く表す互いに無相関ないくつかの新しいパラメータを求め、これをパラメータとする予測モデルによって求めたタイミング点間隔を連結してタイミング点パターンを

生成すること。

(10) 本来定性的パラメータとして表される前後の音韻環境を、音韻の周波数スペクトルを表す種々のパラメータの時間変化という定延的パラメータの形でその他の定量的パラメータに加えて分析することを特徴としたものである。

以下、本発明の実施例に基づいて説明する。

第1図乃至第4図は、それぞれ本発明の第1の実施例を説明するための図で、図中、10はモデル、11は定性パラメータを数量へ変換する変換部、12は加算器、13は減算器で、本実施例においては、多数の音声サンプルを炎め、ラベリングを行って各音韻の時間長を求め、その音韻についての前後音韻環境やモーラ数、モーラ位置などの音韻時間長に影響を与えると考えられる定性的パラメータを求め、これを要因アイテムとして、先に求めた音韻時間長を外的基準にとって数量化I類分析を行い、これによって誤差が最小になるような点線で囲ったモデル10を得、このモデル10によって、制御すべき音韻についての定性パ

ラメータを入力として、最小の誤差の予測値を得る。

第2図は、第1図に示した実施例の変形実施例を説明するための図で、図中、12は加算器、13は演算器、14は係数器で、この実施例は、第1図に示した実施例と同様にして音韻の時間長を求め、その音韻の前後の音韻の時間長やピッチなどの定量パラメータを求め、これによって先に求めた音韻時間長を予測する線形重回帰モデル10を求め、このモデル10によって制御すべき音節についての定量パラメータを入力として最小の誤差の予測値を得るようにしたものである。

第3図は、他の変形実施例を説明するための図で、この実施例は、第2図に示した実施例と同様にして音韻時間長、定量パラメータを得て、このパラメータ群を主成分分析部15によって主成分分析して互いに相関のないデータの特徴をよく表す新パラメータ群に変換し、これを入力定量化パラメータとして以降第2図に示した実施例と同様の処理を行い、これによって最適なパラメータを用

いて制御する。

第4図は、更に他の変形実施例を説明するための図で、図中、16は周波数スペクトルを表わすパラメータ抽出部であるが、まず、第3図に示した実施例と同様にして音韻時間長、定量パラメータを求める。主成分分析は定量的データのみ扱うことが可能であるので、定性的データとしての音韻環境を扱うことはできないが、前後及びそれ自身の音韻、もしくはその調音部分などの周波数スペクトルを表すパラメータ（バンドパスフィルターの出力、LPCケプストラム等）の形で音韻環境を主成分分析の入力として加える。それ以降は、第3図に示した実施例と同様である。これによって音韻の周波数スペクトラム上の特徴及びその時間変化もパラメータとして加えて音韻時間長を制御することが可能である。

第5図乃至第7図は、本発明の他の実施例を説明するための図で、この実施例は、自然なタイミング点パターンを生成するため、タイミング点間隔に影響を与えると考えられる種々のパラメータ

を同時に扱い、多变量統計解析的処理によってパラメータ全体として最適な予測値を得るモデルを作り、これによって予測したタイミング点間隔を連結してタイミングパターンを生成するものである。図中、第1図乃至第3図に示した実施例と同様の作用をする部分には第1図乃至第3図の場合と同一の参照番号を付してある。

而て、第5図に示した実施例は、実音声データを分析し、前後の音韻環境、発話単位のモーラ数、モーラ位置、アクセント情報、強意情報等の定性的パラメータを要因アイテムとし、タイミング点間隔を外的基準に取り、数量化1類分析を行なうことによって実測値との2乗誤差を最小にするような点線で囲った予測モデル10を作るもので、このように構成したモデルによって、求めるべき音節における説明変数を入力としてタイミング点間隔を予測することが出来、このタイミング点間隔を連結してタイミング点パターンが得られる。

第6図に示した実施例は、実音声データを分析

し、前後のタイミング点間隔、ピッチ変化、振幅変化等の定量的パラメータを説明変数とし、タイミング点間隔を目的変数に取り線形重回帰分析を行なうことによって実測値との2乗誤差を最小にするような点線で囲った予測モデルを作るもので、このように構成したモデルによって、求めるべき音節における説明変数を入力としてタイミング点間隔を予測することが出来、このタイミング点間隔を連結してタイミング点パターンが得られる。

第7図に示した実施例は、第6図に示した実施例と同様な定量的パラメータ群を主成分分析によって、互いに相関のない、データの特徴を良く表す新しいパラメータ群に変換し、これを説明変数として以下第6図に示した実施例と同様の処理を行なうことによってタイミング点パターンを得るようにしたものである。

また、対象とするフレームとその前後の適当な数フレームのスペクトルパラメータ（バンドパスフィルターの出力、LPCケプストラム等）をブロック化して扱うことにより、音韻環境の情報も

定量的なデータとして第6図および第7図に示すモデルに取り込み、タイミング点パターンを生成することができる。

上述のように、第5図乃至第7図に示した実施例においては、多変量統計解析法を用いてモデルを構成し、タイミング点パターンを生成することによって自然性の高い合成音声を得ることができるもの。

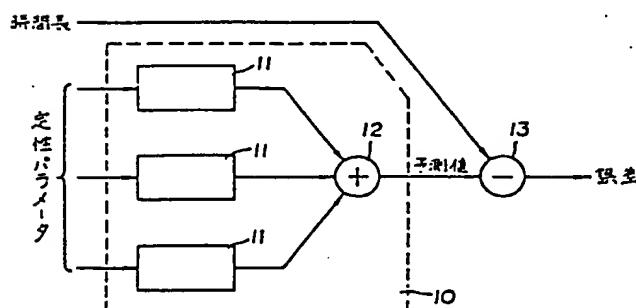
#### 効 索

以上の説明から明らかなように、本発明によると、顎体に影響を与えると考えられるパラメータ全体として、最適なモデルを得ることができ、また、パラメータとしても最も特徴をよく表すよう再構成して制御でき、これによって自然な顎体を生成することができる。

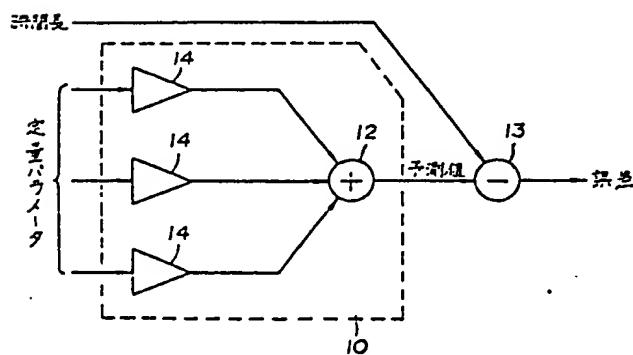
#### 4. 図面の簡単な説明

第1図乃至第4図は、それぞれ本発明の一実施例を説明するための電気的ブロック線図、第5図乃至第7図は、それぞれ本発明の他の実施例を説明するための電気的ブロック線図である。

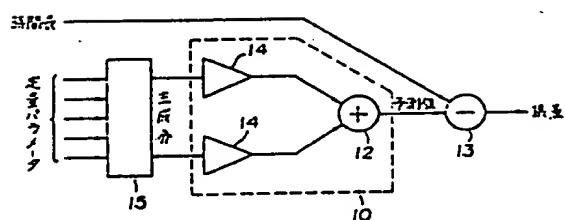
第 1 図



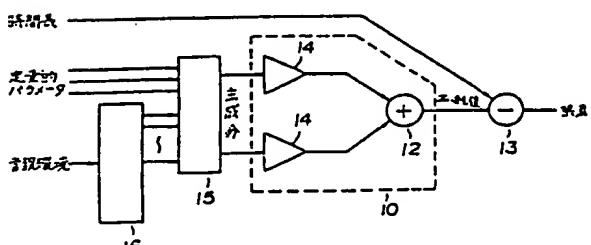
第 2 図



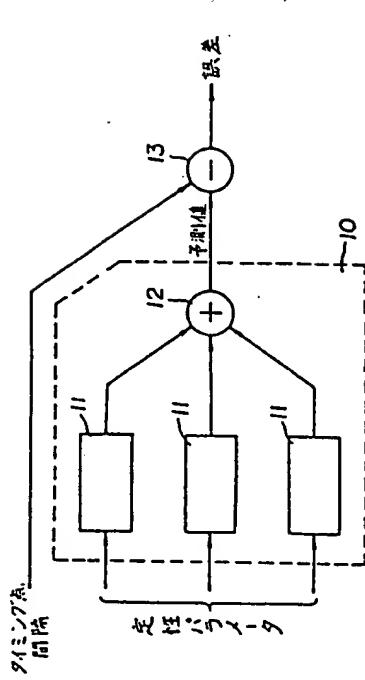
第 3 図



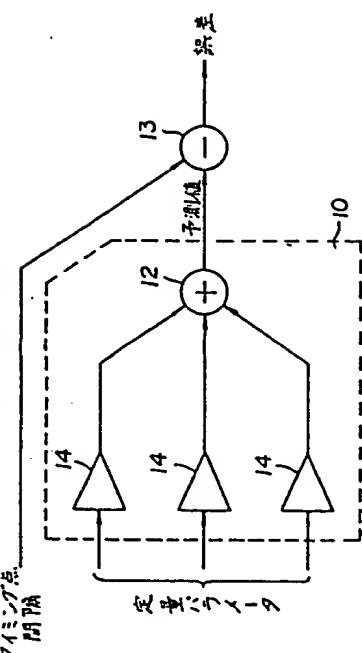
第 4 図



第5図



第6図



第7図

